EXAMINING THE CLINICAL UTILITY AND PREDICTIVE VALIDITY OF DIMENSIONAL MODELS OF PSYCHOPATHOLOGY

Patrick K. Love, M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2021

APPROVED:

Jennifer L. Callahan, Major Professor
Randall J. Cox, Committee Member
Camilo Ruggero, Committee Member
Yolanda Flores Niemann, Interim Chair of the Department of Psychology
Tamara L. Brown, Executive Dean of the College of Liberal Arts and Social Sciences
Victor Prybutok, Dean of the Toulouse Graduate School

Love, Patrick K. *Examining the Clinical Utility and Predictive Validity of Dimensional Models of Psychopathology*. Doctor of Philosophy (Clinical Psychology), August 2021, 48 pp., 4 tables, 4 figures, references, 152 titles.

The *Diagnostic and Statistical Manual of Mental Disorders* arranges co-occurring clusters of symptoms into distinct disorder categories, which theoretically have specific etiologies, pathologies, and treatments. However, researchers and clinicians alike have consistently found *DSM* diagnoses to have high rates of comorbidity, low diagnostic specificity, and no disorder has proven to be a discrete category. There is mounting evidence that dimensional taxonomies more accurately capture the underlying structure of mental illness and clinical presentations. The recently proposed hierarchical taxonomy of psychopathology presumes to address the issues of categorical nosologies using a data driven approach to create a dimensional model of psychopathology. However, heretofore there are no empirical examinations of HiTOP's ability to predict psychotherapy treatment outcomes. This study compared the predictive validity *DSM*, RDoC, and HiTOP criteria using natural language processing on free text narrative notes. Of the three GMM run, only the model using *DSM* criteria as predictors had adequate model fit. Additionally, none of the nosologies significantly predicted treatment course. Implications for the application of RDoC and HiTOP are discussed.

TABLE OF CONTENTS

Page

LIST OF TABLES AND FIGURES

Page

Tables

Figures

CHAPTER 1

CATEGORICAL TAXONOMIES OF PSYCHOPATHOLOGY

Since the introduction of the Feighner criteria (Feighner, Robins, Guze, Woodruff, Winokur, & Munoz, 1972), and the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; APA, 1980), traditional nosologies of mental illness follow the medical model of diagnoses, arranging clusters of symptoms into distinct disorders which are proposed to have specific etiologies, pathologies, and treatments (Zachar & Kenderl, 2007). As a result, clinicians conceptualized the diagnoses listed in the DSM as conditions distinct from one another and distinct from normal functioning (Widiger & Crego, 2018). However, researchers and clinicians alike have noted that the DSM has failed to achieve its goals of accurately modeling the etiologies, pathologies, and effect treatments for the disorders it describes.

Several factors suggest categorical taxonomies misidentify the etiology of psychopathology. First, there is growing evidence that existing diagnoses encompass multiple pathological processes (Clark, Watson, & Reynolds, 1995; Hasler, Drevets, Manji, & Charney, 2004; Meehl & Golden 1982; Skeem, Polascheck, Patrick & Lilenfeld, 2011; Zimmerman et al., 2015). Second, no mental disorder has proven to be a discrete category (Carragher et al., 2014; Haslam, Holland, & Kuppens, 2012; Wright et al., 2013). Third, many clients who present with significant impairment and/or distress fail to meet the diagnostic criteria of any disorder and are often given the Other Specified/Unspecified (formerly Not Otherwise Specified) diagnosis (Kotov et at., 2018). In fact, the Other Specified/Unspecified diagnosis has historically been among the most frequently used in clinical populations (Widger & Edmundson, 2011).

Two common methods have been employed to improve the specificity and accuracy of categorical diagnoses; introducing exclusionary criteria to aid in differential diagnosis, and

splitting disorders using rational divisions. These methods have the goal of creating smaller, more homogenous diagnostic categories while encompassing more diverse clinical presentations (Wideger & Crego, 2018). For example, mixed anxiety-depressive disorder was created to fill the gap between anxiety and mood disorders, schizoaffective between bipolar and schizophrenia, and depressive personality disorder between mood and personality disorders (Frances, 1995). Additionally, binge eating disorder was split out from bulimia and disinhibited social engagement disorder was split from reactive attachment disorder to further improve diagnostic specificity and better capture the complex reality of clinical presentations (APA, 2013). However, these practices have further blurred the boundaries between diagnostic categories (Phillips, Price, Greenberg, & Rasmussen, 2003; Pincus, Frances, Davis First, & Widiger, 1992; Pincus, McQueen, & Elinson, 2003). For example, the DSM-5 Field Trials found 40% of diagnoses did not meet a relaxed cutoff of acceptable interrater reliability (Reiger et al., 2013). Additionally, high rates of diagnostic comorbidity have been consistently found in clinical and research samples when using DSM criteria (Bijl, Ravelli, & van Zessen, 1998; Boyd et al., 1984; Kotov, Perlman, Gamez, & Watson, 2015; Love, 2018; Teesson, Slade, & Mills, 2009; Ormel et al., 2015). These findings suggest that the various iterations of the *DSM* superimposed categorical distinctions onto naturally dimensional phenomena.

The idea that psychopathology is dimensional rather than categorical in nature is not new (Kendell, 1975) and is empirically grounded. The final conference of DSM-5 Research and Planning Work Groups was devoted entirely to implementing dimensional models of substance use disorders, depressive disorders, psychoses, anxiety disorders, developmental psychopathology, and personality disorders (Helzer et al., 2008). For example, there is a large field of research comparing dimensional and categorical conceptualizations of depressive

disorder symptoms. several cross-cultural studies have found a linear relationship between number of depressive symptoms present and level of impairment (Kessler, Zhao, Blazer & Swartz, 1997; Sakashita, Slade & Andrews, 2007; Ustun & Sartorious, 1995). Additional examinations of the structure of major depression have found clinical presentations of major depression shades imperceptibly into "normal", non-clinical experiences of sadness (Andrews et al., 2008). Furthermore, taxometric analyses of self-reported symptoms have found that dimensional models of depression better fit the clinical presentations of adults (Balwind & Shean, 2006; Franklin, Strong, & Greene, 2002; Hankin, Fraley, Lahey, & Waldman, 2005; Ruscio & Ruscio, 2000; Slade & Andrews 2005; Slade 2007) and adolescents (Ambrosini, Bennett, Cleland, & Haslam, 2002; Shields, Giljen, España, and Tackett, 2021). Additionally, dimensional models of psychopathology have shown excellent reliability (First et al., 2002; Hopwood et al., 2017; Kendell, 1989). Despite mounting evidence of the superior performance of dimensional models (Markon, 2010; Kotov, Gamez, et al., 2010; Kotov et al., 2011; Wright et al., 2013) only the dimensional model for Autism spectrum disorders was fully implemented and the dimensional model of personality disorders was relegated to Emerging Measures and Models of the DSM-5 (APA, 2013) leaving these long-standing criticisms unanswered.

CHAPTER 2

ALTERNATIVE DIMENSIONAL MODELS OF HUMAN BEHAVIOR

Research Domain Criteria

Noting the considerable problems with traditional categorical nosologies, two new dimensional models were recently introduced, each using a different theoretical approach to create a more accurate nosology. In 2013, the National Institutes of Mental Health (NIMH) released the research domain criteria (RDoC; NIMH, 2013), a bottom up approach to classifying normal human behavior with the goal of understanding its etiology (Kaufman, Gelernter, Hudziak, Tyrka, & Coplan, 2015). Primarily designed as a research tool, RDoC is organized in a matrix format and is a living document which can change more quickly than the DSM to represent advances in scientific knowledge. The RDoC matrix has eight columns which represent the units used to measure each construct (NIMH, 2021). Current units represented in the matrix include molecules, cells, neural circuits, physiology, behavior, self-reports, and paradigms with neural circuits being the key unit of analyses.

The rows of the matrix represent dimensions of behavior and related constructs are grouped together into higher level constructs called domains. Each domain also contain related lower level constructs. RDoC was originally launched with five domains but a sixth has since been added (NIMH, 2021). Those six domains are as follows. Negative Valence Systems is the domain of neurocircuits responsible for responses to adverse stimuli. This domain contains the constructs of responses to acute threat (fear), responses to potential harm (anxiety), responses to sustained threat, frustrative non-reward, and loss. Positive Valence Systems is the domain primarily responsible for responses to rewarding stimuli and is made up of the constructs reward responsiveness, reward learning, and reward valuation. The Cognitive Systems domain contains

systems responsible various cognitive processes including the constructs of attention, perception, declarative memory, language, cognitive control and working memory. The Systems of Social processes domain contains systems that mediate responses to interpersonal settings and stimuli. It is comprised of the constructs, affiliation and attachment, social communication, perception and understanding of the self, as well as perception and understanding of others. Arousal and Regulatory Systems is the fifth domain of the initial RDoC matrix. It is comprised of the constructs responsible for physiological arousal, circadian rhythms, and sleep-wakefulness. The newest domain, Sensorimotor Systems, is primarily responsible for the control and execution of motor behaviors. This domain is made up of the constructs of motor actions (i.e.: Action planning and selection, sensorimotor dynamics, etc.), Agency and ownership, sensorimotor habit and innate motor patterns.

Critics of RDoC argue that its focus on finding the neurobiological origins of human behavior ignores contextual factors which influence behavior (Lieblich, Castle, & Everall, 2015). This, coupled with RDoC's primary purpose of being a diagnostically agnostic research tool leaves clinicians with no clear distinctions between disordered and normal behavior, limits the clinical utility of the nosology (Lieblich et al, 2015; Wakefield, 2016). However, researchers examining the clinical utility of RDoC have found the dimensional model to more accurately predict the development of mood or disruptive disorders, predicting the length of time before hospital readmission using narrative notes written by RDoC naive clinicians (McCoy et al., 2015), discriminating between subtypes of psychosis (McCoy et al., 2018), predicting symptomology of eating disorders (Wildes & Marcus, 2015), and predicting treatment trajectory of outpatient psychotherapy clients (Love, 2018). While these preliminary findings are very

promising, further research is needed to determine if RDoC accurately captures the structure of human psychology.

Hierarchical Taxonomy of Psychopathology

A second comprehensive dimensional nosology called the hierarchical taxonomy of psychopathology (HiTOP) was proposed by a consortium of clinical researchers with the goal of developing an empirically driven classification system of psychopathology (Kotov et al., 2017). HiTOP emerges from the quantitative classification movement which uses factor analytic approaches to find naturally occurring dimensions among psychopathological symptoms. This approach is unique from the DSM and RDoC which used expert working groups to create the distinctions between categories or domain.

Factor analytic approaches have been used to explain multiple aspects of human behavior. Arguably the best supported of these approaches is the five-factor model of normal personality made up of openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Borkenau & Ostedorf, 1990; Costa & McCrae, 1992; Goldberg, 1990; John, Caspi, Robins, Moffit, & Stouthamer-Loeber, 1994). The strong relationships between these traits and psychological well-being (Krueger, Derringer, Markn, Watson & Skodol, 2012; Malouff, Thorsteinsson, & Schutte, 2005; Trull & Sher, 1994; Widiger & Costa, 2013) coupled with the previously discussed issues with categorical models of psychopathology led to the development of dimensional models of psychopathology using the factor analytic approach. Initial forays into the area found two common factors, internalizing and externalizing behaviors, which predicted the expression of different forms of psychopathology (Lengua, West, & Sandler, 1998; Eisenberg, et al., 2009). For example, externalizing behaviors, such as impulsivity, have been linked to alcohol abuse (Whiteside & Lynam, 2009), substance abuse (Iacono, Malone, and

McGue, 2003), and suicide attempts (Verona, Sachs-Ericsson, & Joiner, 2004). Similarly, internalizing behaviors in children predict the development of social withdrawl, social skill deficits sometimes associated with developing depression (Rubin & Mills, 1991). The internalizing and externalizing factors have been found in child (Blanco et al., 2015; Martel et al., 2017; Waldman et al., 2016), adolescent (Achenbach et al., 1991; Beauchine & Hinshaw, 2017; Waldmen et al., 2016), adult (Carragher et al., 2014) and cross-cultural samples (Krueger, Chentsova-Dutton, Markon, Goldberg, & Ormel, 2003). When examined together, internalizing factors were found to explain much of the comorbidity between depression, anxiety, post-traumatic stress disorders; while alcohol and substance abuse, attention deficit-hyperactivity, oppositional defiant, and antisocial-personality disorders (Blanco et al., 2015; Krueger et al., 2003; Lilienfeld, 2003; Mesman, Bongers, & Koot, 2001; Whiteside & Lynam, 2009). These studies revealed a higher-order dimension of psychopathology called the General P factor of psychopathology (Caspi et al., 2014). In their review, Forbes et al. (2016) introduced a hierarchical model of internalizing/externalizing spectra which was phenotypically stable, linked to genetic vulnerabilities, applicable developmentally, and predictive of different problematic behavior developing during childhood (i.e., withdrawn, anxious/depressed, attention problems, and/or aggressive behaviors; Achenbach, 2009; Lahey et al., 2017). However, this model failed to account for psychotic disorders and personality psychopathology.

Categorical diagnoses of personality disorders have been some of the most problematic diagnostic categories for clinicians and researchers (First et al., 2002; Kendell, 1989). Like other DSM diagnoses, comorbidity among personality disorders is the norm (Anderson et al., 2014; Clark, 2007; Lilienfield et al., 1994; Livesley, 2003) which suggest that arbitrary boundaries have been placed onto dimensional phenomena. In contrast, there is an abundance of empirical

support for dimensional models of personality psychopathology (Fowler et al., 2015; Krueger & Markon, 2014; Markon & Krueger, Markon, Patrick, &v Iacono 2005; Morey et al., 2007; Rodriguez-Seijas, Eaton, & Krueger, 2015; Wygant et al., 2016). Additionally, there is no evidence empirical evidence that personality pathology is categorical in nature (Clark, 2007; Trull & Durrett, 2005) or are distinct phenomena from other major mental illness diagnoses in the DSM (Krueger, 2005).

Personality has been linked to psychopathology (Kotov et al., 2010; Samuel & Widiger, 2008; Widiger & Tull, 2007) primarily using the Five Factor Model of personality (FFM): openness-to-experience (O), conscientiousness (C), extraversion (E), agreeableness (A), and neuroticism (N; Costa & McCrae, 1992). For example, consistent relationships have been found between high N and depression and anxiety disorders (Kotov et al. 2010); high C and anorexia (Cassin & von Ranson, 2005); and high N, low A, and C with many personality disorders (Widiger & Mullins-Sweatt, 2009). While there have been some studies linking psychotic disorders with the FFM (Ross, Lutz & Bailey, 2002; Tackett et al., 2008; Taylor & Bagby, 2012; Uliaszek, Al-Dajani, & Bagby, 2015). Thus, the alternative dimensional model of personality psychopathology included in the DSM-5 will likely prove to be insufficient at addressing issues of comorbidity in psychiatric diagnoses.

The unified, quantitative approach used to develop HiTOP aims to rectify the problems found in traditional categorical models of psychopathology and the artificial division between so called Axis I and Axis II disorders (Kotov et al., 2017). HiTOP aims to address these issues in several ways. First, HiTOP constructed dimensional syndromes based on the empirical covariation of symptoms and only groups related symptoms together thereby creating unitary constructs (Clark & Watson, 2006) instead of relying on arbitrary thresholds between syndromes

created using expert consensus which have proven to be quite heterogeneous. This method also improves upon the reliability of categorical taxonomies as evidenced by the high test-retest reliability of dimensional constructs (Watson, 2003). HiTOP's hierarchical structure in which related syndromes are subsumed under higher order spectra address comorbidity commonly found in clinical presentations (Kotov at al., 2017). While, high rates of comorbidity indicates problems with the categorical structure of the DSM, it also suggest that psychopathology develops along common pathways which are determined by shared pathways, pathological process and predicts the courses of illness (Brown & Barlow, 2009; Krueger & Markon, 2011). For example, comorbid presentations commonly occur between clients with anxiety and depression because the individual has a predilection toward internalizing behaviors with individual differences determined by various environmental factors. Thus, the hierarchical organization of the dimensions of HiTOP can accurately account for the presentation, severity, and course of individual patients (Kotov et al., 2017).

Finally, HiTOP's dimensional structure allows all clients to be described using the taxonomy because all levels and types of pathology can be described using the same language (Kotov et al. 2017). Joint structural analyses of DSM clinical and personality disorders yield six common spectra. Kotov, Ruggero and colleagues (2011) analysis of 2,900 outpatients with 25 different disorders found recognizable dimensions of internalizing comprised of anxiety, depressive, dependent, obsessive-compulsive, borderline, and paranoid PD; externalizing comprised of substance use, antisocial behavior and conduct problems; thought disorder comprised of psychotic disorders, bipolar I, schizotypal, paranoid, and schizoid PD; antagonism comprised of histrionic, narcissistic, borderline, and paranoid PD; and somatoform which included undifferentiated somatoform disorder, hypochondriasis, and pain disorder. While they

were the first to identify a sperate somatoform factor, a previous study by Røysamb et al. (2011) also identified internalizing, externalizing, antagonism factors, and a pathological introversion/detachment factor. These differences are likely because Røysamb et al. (2011) sample did not include psychotic disorders or mania which were present in Kotov, Ruggero et al.'s (2011) sample. Two other studies that operationalized psychopathology based off heterogenous samples identified six spectra which mirrored those found in the earlier studies: negative affectivity (internalizing), psychoticism (thought disorder), disconstraint (externalizing), aggressiveness (antagonism), introversion (detachment), and somatization (somatoform) (McNulty & Overstreet, 2014; Sellbum, 2017).

Building upon the exiting literature exploring dimensional models of psychopathology, six spectra were included in the initial HiTOP model. Those six spectra include internalizing (or negative affectivity), thought disorder (or psychoticism), disinhibited externalizing, antagonistic externalizing, detachment, and somatoform (see Figure 1 for the complete HiTOP model proposed by Kotov et al., 2017). Additionally, 7 subfactors (Sexual problems, eating pathology, fear, distress, mania, substance abuse, and antisocial behavior) were proposed sitting between the higher order *p*-factor of psychopathology and lower order empirical syndromes, and symptoms components and maladaptive traits. Since its original proposal in 2017, the overall dimensional structure of HiTOP has been argued for by 10 meta-analytic reviews examining more than 1,100 studies published by members of the HiTOP consortium (Conway et al., 2019; Hopwood et al., 2020; Kotov et al., 2020; Kotov et al. 2021; Krueger et al., 2018; Latzman et al. 2020; Perkins et al., 2020; Ruggero et al. 2019). Additionally, Perkins et al. (2020) and Kotove et al. (2021) have provided tentative evidence in support of three superspectra; emotional dysfunction, psychosis

and externalizing. These superspectra were excluded from this analysis as they are still provisional in nature (Kotov et al., 2021).

Although a complete explanation of the validation of these spectra is beyond the scope of this project (see Kotov et al., (2017) and Kotov et al. (2021) for the complete validation procedures), it is important to mention several things about the current HiTOP model. First, the lower order disorder/syndrome level of the model is described in terms of DSM-5 diagnoses to assist individuals unfamiliar with the model understand what types of behaviors are included in each spectra. Second, neurocognitive and neurodevelopmental clusters have thus far been excluded from structural studies and thus were not included in the current iteration of HiTOP even though they are included in the DSM-5; and thus, the current HiTOP model is not yet comprehensive even though it does include the majority of psychopathology (Kotov et al., 2017; Kotov et al. 2021). Third, the current HiTOP model does not yet integrate features of the progression of illnesses as there is a general lack of longitudinal research data (Kotov et al. 2021). Fourth, HiTOP was constructed using data collected primarily from adolescents and adults up to age 65 (Kotov et al., 2017), which means it is unclear what the structure of psychopathology is in children below age fifteen and adults over 65. The little research that does exist supports the exitance of the general *p*-factor of psychopathology (Kotov et al., 2021; McElroy et al., 2018), but the exact structure of the lower order dimensions remains unclear.

Hypotheses

While the goals of the new RDoC and HiTOP taxonomies show promise in achieving their goals of more accurately describing the human psychological experience, it is yet unclear if they will prove to be more useful to clinicians in real world applications. There is emerging evidence that RDoC can accurately predict treatment outcomes (Love, 2018: McCoy et al., 2015;

McCoy et al., 2018; Wildes & Marcus, 2015). Additionally, several articles have proposed how clinicians can implement the HiTOP framework to conceptualize their clients (Eubanks & Hunter, 2020; Hopewood et al., 2020; Ruggero et al., 2019) and integrate the model into clinical training (Ruggero, Callahan, et al., 2019). Previous research has also shown the utility of individual HiTOP constructs to predict suicidality, and future-treatment seeking among others (Morey et al. 2012; Sunderland & Slade, 2015). However, no study has been published which examines the predictive validity of the entire HiTOP framework for psychotherapy treatment outcomes. Furthermore, it is unclear how much difficulty current clinician's will have adopting and applying a HiTOP's hierarchical dimensional nosology to their conceptualizations of clients in real world settings.

Thus, this article aims to explore these two topics by utilizing archival records from an outpatient psychology training clinic. First, natural language processing will be used to determine how well HiTOP's language matches the language currently used by clinicians. Second, HiTOP's ability to predict treatment psychotherapy outcome will be examined and compared to that of DSM diagnostic criteria. Finally, the predicative validity and fit of HiTOP criteria to the natural language of clinicians will be compared to that of RDoC to directly compare the strengths and weakness of the three nosologies.

CHAPTER 3

METHOD

Participants

This project used archival data from adults who sought individual psychotherapy treatment at the University of North Texas (UNT) Psychology Clinic over a 5-year period (2011-2015. The UNT Psychology Clinic is a training clinic for students in APA accredited clinical, counseling, and clinical health psychology doctoral programs at UNT. All services were provided by closely supervised trainee clinicians affiliated with these programs. Additionally, all clients are informed of the research mission of the training clinic during their first session. All clients of the UNT Psychology Clinic were asked to complete a standardized demographic form as part of their intake paperwork. The form collects a broad range of demographic information including age, self-identified race/ethnicity, gender, marital status, religion, occupation, income, high school graduation, and veteran status. Only clients who consented to the use of their de-identified treatment records are included in the sample. Archival data was accessed following procedures approved by the UNT Institutional Review Board (IRB) and the Psychology Clinic Executive Committee (PCEC).

Inclusion / Exclusion Criteria

Consented clients included in this project have (1) completed the Psychiatric Diagnostic Screening Questionnaire before their intake session, (2) attended at least two sessions of individual psychotherapy, and (3) completed the Outcome Questionnaire – 45 (OQ) before each psychotherapy session. In addition, intake summaries and termination reports written by the clinician were present in the record. The initial sample consisted of 1,000 clients who attended at least two sessions of psychotherapy and had both an intake and termination report in their file.

Of those, 356 clients were removed from the sample for not having a PDSQ which was completed at intake. A further 322 clients were removed from the sample for having incomplete OQ data. Thus, 322 clients met inclusion criteria and were included in the final sample. Clients in this sample completed a maximum of 111 outpatient psychotherapy sessions, a mean of 16 outpatient sessions and a median of 11 sessions. The demographic information of the final sample of 322 participants used in the models is included in Table 1.

Measures

Outcome Questionnaire 45.2 (OQ)

The OQ is a 45 item self-report measure designed to assess the level of psychological distress a client experienced during the previous week (Lambert, et al., 2004). The OQ is one of the most commonly used measures used to monitor psychotherapy outcomes over time (Vermeersch et al., 2004) and demonstrates sensitivity to change during treatment, high specificity to change, and high concurrent validity (Vermeersch, Lambert, & Burlingame, 2010). To measure psychological distress, clients are asked to respond to how frequently they experienced the listed symptoms using a 5-point Likert-scale. The scale is anchored with responses ranging from *almost never* (0) to *always* (4). The OQ item responses are then summed to yield a total score ranging from 0 to 180. According to the manual, total raw scores of 63 and above indicate clinically significant levels of distress. A change of 14 points or more indicates a reliable change in psychological distress between sessions. A reduction of total OQ score equal to or greater than 14 points from the time of intake indicates a reliable improvement in symptom severity while a 14-point or greater increase indicates reliable deterioration and increase of psychological distress. Lambert et al. (2004) recommend that missing data be substituted with the mean of the subscale that question loads onto. However, independent research has failed

found high correlation between the OQ subscales indicating they lack validity (Boswell et al., 2013; Umphress et al., 1997). Therefore, multiple imputation with all present items will be used when OQ has less than 10% of missing data points, as this method creates the least amount of bias in the data (Biering, Hjollund, & Frydenber, 2015). OQ's missing answers to 10% or more questions were excluded from the analyses. Thirtysix clients' data were removed from the sample for this reason.

Psychiatric Diagnostic Screening Questionnaire (PDSQ)

The PDSQ is a self-report measure designed to identify respondents who are likely to meet criteria for 13 psychological disorders associated with the Axis I disorder in the DSM-IV TR (Zimmerman and Mattia, 2001). The measure contains 111 dichotomous ("yes" or "no") items asking the respondent to indicate which psychological symptoms they had experienced during the previous two weeks or six months, depending on the item. The measure yields a clinical disorder subscale for each of the 13 clinical disorders which include 90% sensitivity clinical cut off scores and critical items for each clinical disorder. The PDSQ subscales have a mean negative predictive validity of 97% and demonstrate good to excellent levels of internal consistency, test-retest reliability, and convergent validity (Zimmerman and Mattia, 2001). See Table 2 for the frequency of clients who met or exceeded each PDSQ clinical cut score.

Assembling, Cleaning, and Tokenizing the Corpus

The free text narratives of intake reports, termination reports and progress notes associated with all clients in the initial sample (n = 1,000) were combined to create the corpus for the analysis. This larger corpus of text was to create each bin from a diverse author set, thereby reducing the influence of writing style on keyword selection. Next, the free text was cleaned using the natural language processing methods outlined by Kao and Poteet (2007), and Bird,

Klien, and Loper (2010) using a Unix command line to execute the commands. The specific process was carried out as follows. First, all text was converted into lower case. Second, all abbreviations, contractions, numerical text, ordinal text, and symbols were converted into their long-form equivalents (i.e.: 1 into one; 1st into first, $ into dollar, etc.). Third, all extra white space, and punctuation were removed from the corpus. Fourth, all commonly used, but semantically unmeaningful words, called stop words, were removed from the corpus. There are several commonly used stop word lists used for natural language processing, with each list having their own strengths and drawbacks. The stop word list "stopwords en" from commonly used R text mining package "TM" was used as it is the least restrictive list and allowed for the most flexibility when creating the estimated RDoC and estimated HiTOP scores (Lewis et al., 2004). Fifth, the text was stemmed to truncate the words to their root form using SnoballC text stemming package (Tomar, 2017). Finally, cleaned text was tokenized into unigrams, bigrams, and trigrams (one word, two-word pairs, and three-word triplets) before being processed to estimate RDoC (eRDoC) and HiTOP (eHiTOP) scores. Calculation of these estimated scores (i.e., eRDoC and eHiTOP) is described in Chapter 4.

CHAPTER 4

RESULTS

After the free text narrative notes were cleaned and tokenized using the above methods, the unigrams, bigrams, and trigrams were sorted into bins reflecting the six RDoC domains and six HiTOP spectra. To do this, two expert coders sorted the terms independently first into bins representing the six RDoC domains. Then, they independently sorted the clean tokenized term list as second time into bins representing the six HiTOP spectra. The spectra level of the HiTOP model was chosen as it would provide the most analogous comparison with RDoC domains. Coders were trained to complete this task by reading seminal articles on HiTOP's structure (e.g. Kotov et al. 2017, etc.), reviewing the content in the RDoC matrix, and discussing what types of behaviors are representative of the HiTOP spectra and RDoC domain. Coders were not blind to the purpose of the study nor its hypotheses. After the initial coding was completed, the coders met and discussed their coding efforts to come to 100 percent agreement about the structure of the bins. Next, the cleaned and tokenized narrative texts from each of the 322 intakes included in the final sample were compared to the key terms in each RDoC and HiTOP bin; finding the percent fit (frequency of key words appearing in the narrative note divided by the total number of words in the cleaned and tokenized intake report) to create eRDoC scores and eHiTOP scores for each domain and spectra. Thus, each intake report was assigned an eRDOC score for each RDoC domain and an eHiTOP score for each HiTOP spectra. Upon examination of the data, both eRDoC and eHiTOP scores were found to be significantly skewed. A square root transformation was performed to correct this skewness while appropriately handling some true zero values found in the eHiTOP and eRDoC scores.

Three separate growth mixture models (GMM) with linear trajectories were

independently estimated. To tolerate the variable time between OQ measurement points, a structural equation model (SEM) framework was used for each GMM. All GMMs were conducted using the Mplus version 8.5 software (Muthén, & Shedden, 1999; Muthén, et. al, 2002; Muthén, 2001a; Muthén, 2001b; Muthén, 2004; Muthén & Masyn, 2005) The first GMM used PDSQ cutoff scores as time invariant predictors of OQ score over time. The second and third GMM models were conducted using the eRDoC domain scores and eHiTOP spectra scores respectively as time invariant predictors of OQ score over time. Treatment course was measured up-to the sample mean of 16 treatment sessions, as opposed to an unconstrained number of sessions, to reduce problems with model convergence. Missing data due to treatment attrition was handled using robust FIML estimation due to its use of all available data points, simultaneously accounting for data with non-normal distributions, and accounting for these differences with adjustments to standard errors and scaling the chi-squared statistics (Little et al., 2014).

Fit of the three GMMs was determined using the following conventions. Good model fit is said to occur when the confirmatory fit index (CFI) and Tucker-Lewis index (TLI) are ≥ .90.; when the standardized root mean square residual (SRMR) is ≤ .08; when the root mean square error of approximation (RMSEA) is ≤ .08; and the value of RMSEA is within 90% confidence intervals (Weston & Gore, 2006). Chi-squared fit measures were divided by the models' degrees of freedom to correct for biases caused by sample size. Standard conventions state that corrected chi-square ratios falling below 3 are indicative of adequate model fit and, corrected ratios falling below 2 are indicative of good model fit (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006). To compare fit between models their Akaike information criteria (AIC) and Bayesian information criteria (BIC) statistics were reported, with smaller AIC and BIC values

indicating better fitting models (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006).

See Table 3 for each of the models' fit statistics. The overall fit of the final PDSQ model was adequate with the chi-square ration, RMSEA 95% confidence interval and the SRMR statistics falling in the good range. However, the TLI fell in the adequate fit range and the CLI indicated poor model fit. The fit of the both the eRDoC model and eHiTOP spectra model were marginal with each models' RMSEA, SRMR, and TLI statistics meeting the adequate level. To improve eRDoC model fit and reduce the chance of overspecification, the newest and least validated RDoC Sensory Motor domain was removed from the model. Similarly, a second HiTOP model was created by sorting tokenized terms into bins representing eHiTOP subfactors following the previously described procedure. However, the eRDoC five domain and eHiTOP subfactor models did not fit the data better than the original eRDoC domain and eHiTOP spectra models, and thus were discarded. See Table 4 for a comparison of different eRDoC and eHiTOP fit statistics.

When comparing the fit between models, the PDSQ GMM fit these data the best statistically, with the eRDoC Domain GMM and eHiTOP Spectra GMM fitting the data equally well. Given the adequate fit of the PDSQ GMM further examination of the model's factor loadings are warranted. In the PDSQ GMM model, the cut scores for major depressive disorder, psychosis, substance use disorder, generalized anxiety disorder, somatoform disorder, and hypochondriasis had significant positive intercept loadings, meaning they were associated with higher OQ scores before the initial treatment session. No PDSQ cutoff score significantly predicted the trajectory of OQ scores over the course of treatment. See Figure 2 for PDSQ GMM factor model and factor loadings.

Given the marginal fit of the eRDoC Domain GMM and eHiTOP Spectra GMM, those model's factor loadings are interpreted with caution. The eRDoC negative valence domain score had a significant intercept loading suggesting that it was associated with higher OQ scores at the initial treatment session. There were no other significant intercepts in the eRDoC Domain GMM. Additionally, the cognitive domain eRDoC scores had a significantly negative slope, thus it predicted a significant decrease in OQ score over the course of treatment. Qualitatively, examination revealed the most frequently used words in the cognitive domain bin include dwell, depression, mania, forgetfulness, concentration, impulsive, irrational, and repressed. This suggests that clinicians who noted these symptoms at intake had clients who were more likely to improve over the course of psychotherapy. There were no other significant predictors of treatment trajectory in the eRDoC domain GMM. See Figure 3 for the eRDoC Domain GMM factor model and factor loadings.

Like the PDSQ model, the eHiTOP Spectra GMM, none of the eHiTOP spectra predicted psychotherapy treatment trajectory. However, the eHiTOP internalizing spectra was associated with significantly higher OQ scores at the initial treatment session. See Figure 4 for the eHiTOP Spectra GMM factor model and factor loadings. Correlations between eHiTOP spectra scores and OQ scores were conducted to determine the relationship between self-reported and clinician reported psychological symptoms. A Bonferroni correction was used to control for type one error across the multiple analyses. There were no statically significant correlations between eHiTOP spectra scores and OQ scores across the 16 treatment sessions. This suggests that there is not a strong relationship between eHiTOP scores and client-reported psychological distress as measured by the OQ-45. To ensure the poor eRDoC and eHiTOP model fit were not due to coding biases, an independent auditor examined 10% of the eRDoC and eHiTOP bins. After

review, the auditor's ratings were in 74% agreement with the sorting of the eRDoC bins and 67% agreement of the sorting in eHiTOP bins, suggesting there is likely bias that contributed to the poor fit of eRDoC, eHiTOP GMMs and the poor correlations between eHiTOP spectra scores and OQ scores.

CHAPTER 5

DISCUSSION

This study had two main aims. First, to determine how well the HiTOP model maps onto clinical conceptualizations in a naturalistic setting. Second, to compare the predictive utility of the DSM, HiTOP and RDoC nosologies with respect to psychotherapy outcomes among a sample of outpatient psychotherapy clients. Based on a rapidly growing body of literature, it was hypothesized that the eHiTOP spectra scores would have equal or greater predictive validity than eRDoC domain scores, and both dimensional models would have greater predictive validity than DSM criteria. The results of this study were quite surprising and did not match the original hypotheses. Of the three GMM models conducted, only the model using DSM criteria, represented by PDSQ cut scores, adequately fit the data set. More specifically, clients who met cut scores for major depressive disorder, psychosis, substance use disorder, generalized anxiety disorder, somatoform disorder, and hypochondriasis reported significantly more distress before their initial psychotherapy session than clients who did not meet those cut scores. However, PDSQ cut scores did not significantly predict treatment outcome. This lack of predictive validity is unsurprising given the extremely high rate of comorbid DSM diagnoses (87.3%) found in this sample. These findings are similar to those found by Love (2018) suggesting that despite the many, and well, documented shortcomings of the DSM, its categorical framework holds capacity for identifying experiences which both clinicians and clients identify as impairing in some way.

In contrast, both the eRDoC domain score GMM and eHiTOP spectra score GMM had poor model fit. Thus, even though the negative valiance eRDoC domain and internalizing spectra eHiTOP scores were associated with higher OQ scores before initial treatment, and the cognitive systems eRDoC domain scores predicted significant reductions in symptom distress over the

course of treatment, it is unclear if these findings are meaningful. However, when you also consider that eHiTOP categories were not significantly correlated with self-reported psychological distress suggest, it is unlikely that the significant factor loadings in the RDoC and HiTOP models are statistically meaningful. Therefore, it would be ill advised to put too much stock into any interpretations about the clinical utility of either RDoC or HiTOP nosologies based on these findings.

This poor fit was likely due to multiple factors. First and foremost, the subjective nature of the coding procedure likely led to imprecise key words. Second, the decision to restrict key token coding to one bin per nosology likely impacted the fit of the eHiTOP model. While this was originally done to reduce the chance of collinearity within eHiTOP and eRDoC scores, this does not align with the spirit of HiTOP which allows spectra to be correlated.

Potential Implications for Application of Dimensional Nosologies and Future Directions

As previously stated, it is difficult to make strong conclusions based on the current models due to the biases present in the coding system for eRDoC and eHiTOP scores. However, there are a few possible implications that can be gleaned from the results of the study. First, the lack of predictive validity of the categorical DSM criteria and high rate of comorbidity support the argument that categorical nosologies do not accurately reflect the natural organization of psychopathology (Carragher et al., 2014; Haslam, Holland, & Kuppens, 2012; Kotov et al. 2018; Kotov et al. 2021; Wright et al., 2013). Additionally, this muddy clinical picture makes it extremely difficult for clinicians to identify the primary presenting complaints and then select the treatment option which is most likely to be effective. The results of this study cannot directly support how dimensional models like HiTOP and RDoC solve this issue, it does not contradict the arguments of their proponents either.

Second, there is a large body of research detailing how clinician ratings and patient self-report ratings can reveal related but distinct clinical presentations (e.g.: Corruble et al., 1999; Grey & Kennedy, 1993; Uher et al., 2012). In this study, both clinician ratings in the form of free text narrative notes, and client self-report measures (PDSQ and OQ) were used. This could have contributed to the difference in fit between the three models. The PDSQ GMM model, which relied self-reported data for both the predictor variables and outcome variables, was the only model which adequately fit the data. In contrast, the eRDoC and eHiTOP GMM models used predictors derived from narrative notes written by trainee clinicians. The difference in clinician identified and client identified symptoms could have further contributed to the poor fit of the eRDoC and eHiTOP models in this study.

An important limitation of this sample, albeit likely representative of the current state of the field, is that study clinicians were likely relying on DSM criteria to conceptualize and write about the clients. Indeed, the intake reports used in this study were written before either RDoC or HiTOP had been formally proposed and disseminated to the public. One explanation for this is that cognitive bias of the clinician likely also contributed to the poor fit of the eRDoC and eHiTOP models (Conway & Brown, 2018). Using training methods outlined by the Clinical Translation Workgroup of the HiTOP consortium could help remedy part of these issues by teaching clinicians how to apply dimensional nosologies to their clinical conceptualizations (Eubanks & Hunter, 2020; Hopewood et al., 2020; Ruggero et al. 2019; Ruggero, Callahan, et al. 2019). It is possible that after clinicians are trained in HiTOP case conceptualization and treatment planning, the language used in their intake reports would have more predictive validity for psychotherapy treatment outcome.

An exploration of previous research seeking to validate the clinical utility of the entire

RDoC and HiTOP models indicate a potential gap in the HiTOP research base. Previous research on the clinician utility of RDoC using natural language processing on data sets with the same clinician biases as the present study not only found GMM and other statistical models that fit the data, but also found that RDoC had higher predictive validity of treatment outcomes than DSM criteria in various patient populations (i.e., Love, 2018; McCoy et al., 2105; McCoy et al., 2018). These findings indicate that it is possible for to overcome clinician biases in training and derive meaningful results using natural language processing techniques on dimensional models.

It is unreasonable and unhelpful to expect clinician and client conceptualizations of their functioning to match perfectly. Even so, there needs to be a basic level of agreement between what the client and clinician identify as problematic, initial treatment goals, and the intervention used to achieve those goals for psychological intervention to be effective. The well-established literature exploring the effectiveness of collaborative treatment planning underscore this fact. Among its many benefits, collaborative treatment planning has been shown to increase treatment adherence (Stanhope, Ingoglia, Schmelter, & Marcus, 2013), reduce the frequency of suicidal ideation and suicidal acts (Jobes, 2012), and improve the therapeutic alliance (Chinman et al., 1999). In fact, collaborative treatment planning is such a foundational part of most evidenced based psychotherapy treatments, it is included in many states' licensure and provider guidelines (e.g. Commonwealth of Massachusetts 130 CMR 429.000; SAMHSA, 2015). Thus, it is encouraging to see the Clinical Utility Working group of HiTOP actively engaged in training efforts and field trials improve clinician understanding of this model (Eubanks & Hunter, 2020; Hopewood et al., 2020; Ruggero et al. 2019; Ruggero, Callahan, et al. 2019). This training is important considering that Conway and Brown (2018) conducted a study attempting to examine if the latent structure of HiTOP could be found in their large sample of outpatients. More

specifically, they used a confirmatory factor analysis to determine if the internalizing and externalizing factors were present in data collected via a well validated structured clinician administered interview. They too found poor model fit and suggested it could have been caused by clinician cognitive bias by the lack of training in HiTOPian case conceptualization. The Clinical Utility Working group's emphasis on clinician training should solve this issue.

Third, the fact that neither RDoC domains nor HiTOP spectra fit these data could indicate that both models maybe incomplete. In fact, the creators of both nosologies openly admit they are both works of progress and have changed over time to incorporate immerging data (Kaufman et al., 2015; Kotov, et al. 2017; Kotov et al. 2021; NIMH, 2013; NIMH, 2021). Several authors, including the HiTOP Consortium itself, argue that RDoC domains and HiTOP can be used to identify the etiology of maladaptive behaviors. Although directly assessing the relationship between RDoC and HiTOP criteria was outside of the scope of this project, the results offer a tentative look into the complimentary nature of RDoC Domains and HiTOP Spectra. For example, the possible positive correlation between negative valence eRDoC domain scores and the internalizing eHiTOP spectrum scores with OQ scores at the start of psychotherapy could suggest those constructs account for psychological distress experienced by patients. Additionally, the tentative finding that only cognitive eRDoC domains scores predicted treatment trajectory could suggest that conceptualizing cases with the RDOC frame is likely beneficial for clinicians. Research into the biological correlates of potential threat linked stronger startle potentiation with increased likelihood of having a disorder related to the fear subfactor (e.g., social anxiety) as opposed to distress related disorders (e.g. generalized anxiety and depression) (Gorka et al., 2017). See Michelini et al. (2020) and Kotov et al. (2021) for more comprehensive reviews of these associations. Thus, additional research should be conducted which directly examines the

relationship between RDoC and HiTOP nosologies and their combined clinical utility. It is also important to note the sample used in this study were predominantly American, White, cisgender well educated, single individuals of college age of unknown sexual orientation. This limits the generalizability of the study to non-majority populations including BIPOC and members of the LGBTQIA+ communities and those of varying educational levels. Future research should use samples which overrepresent these communities to ensure that the structure of these nosologies holds across ethnic and cultural backgrounds.

Although this study failed to identify support for the predictive validity of either the HiTOP or RDoC nosologies, it was not without its own utility. First, it showed that DSM-5 criteria do effectively capture behaviors which cause individuals to feel distress, despite the numerous issues associated with that nosology. This is an important baseline to note and with which to compare all subsequently proposed models of psychopathology. Second, it provided tentative support of the argument that RDoC and HiTOP are complimentary nosologies that, taken together, might improve upon the DSM-5 (Michelini et al., 2020; & Kotov et al. 2021). Future research should explore how the dimensional spectra of HiTOP interface with RDoC domains to illuminate not only the natural structure of psychopathology, but also, its etiology (Perkins et al., 2020). Doing so could improve the efficacy of mental health treatment and help create the personalized treatment plans which the creators of HiTOP (Kotov et al. 2017), and RDoC (Insel, 2014) both tout as the future of mental health treatment.

Table 1

Demographics

Variable		Overall Sample
Age, M (SD)		27.1 (10.96)
Household Family Income, M (SD)		$13,070 ($39,476.65)
Gender, n (%)	Male	106 (32.9%)
	Female	223 (65.8%)
	Transgender	3 (0.7%)
	Missing	2 (0.6%)
Ethnicity, n (%)	White, Non-Hispanic	210 (65.2%)
	African-American	18 (5.6%)
	Latino(a)/Hispanic	35 (10.9%)
	Asian/Pacific Islander	12 (3.7%)
	Middle Eastern/Southeast Asian	7 (2.2%)
	Native American/Alaskan Native	5 (1.6%)
	Multiracial	23 (7.8%)
	Other/Did not identify	10 (3%)
Marital Status, n (%)	Single	258 (80.1%)
	Married	35 (10.9%)
	Divorced	15 (4.7%)
	Other	12 (3.7%)
	Missing	2 (0.6%)
High School Diploma, n (%)	Yes	304 (94.4%)
	No	14 (4.3%)
	Missing	4 (1.2%)
Veteran Status, n (%)	Yes	13 (4.1%)
	No	305 (94.7%)
	Missing	4 (1.2%)

Table 2

*Number of Clients Who Met or Surpassed the PDSQ Cutoff Criteria by Diagnostic Category (*N *= 322)*

Disorder	Number of Clients Who Meet Cutoff Score	Percent of Sample
MDD	203	63%
PTSD	126	39.1%
Bulimia/Binge Eating	39	12.1%
OCD	114	35.4%
Panic Disorder	131	40.7%
Psychosis	65	20.2%
Agoraphobia	130	40.4%
Social Phobia	219	68%
Alcohol Abuse/Dependence	81	25.2%
Drug Abuse/Dependence	54	16.8%
GAD	192	59.6%
Somatization Disorder	105	32.6%
Hypochondriasis	78	24.2%
Met No Cutoff Value	11	3.4%
Met One Cutoff Value	30	9.3%
Met Multiple Cutoff Values	281	87.3%

Table 3

Fit Statistics of PDSQ, eHiTOP and eRDoC GMM Models

Fit Statistic	PDSQ GMM	eHiTOP Spectra GMM	eRDoC 6 Domain GMM
χ^2, (df; p-value)	598.66 (313; 0.000)	480.344 (215; 0.000)	503.18 (215; 0.000)
χ^2/df ratio	1.91**	2.23	2.34
RMSEA (95% CI)	0.053 (0.047 – 0.060)**	0.062 (0.055 - 0.069)*	0.065 (0.057 - 0.072)*
SRMR	0.05**	0.07*	0.065*
CFI	0.94	0.935	0.932
TLI	0.93*	0.94*	0.932*
AIC	26044.43	7769.73	7764.34
BIC	26221.84	7894.293	7888.90

Note: * Indicates adequate fit according to rules of thumb (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006). ** Indicates good fit according to rules of thumb (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006).

Table 4

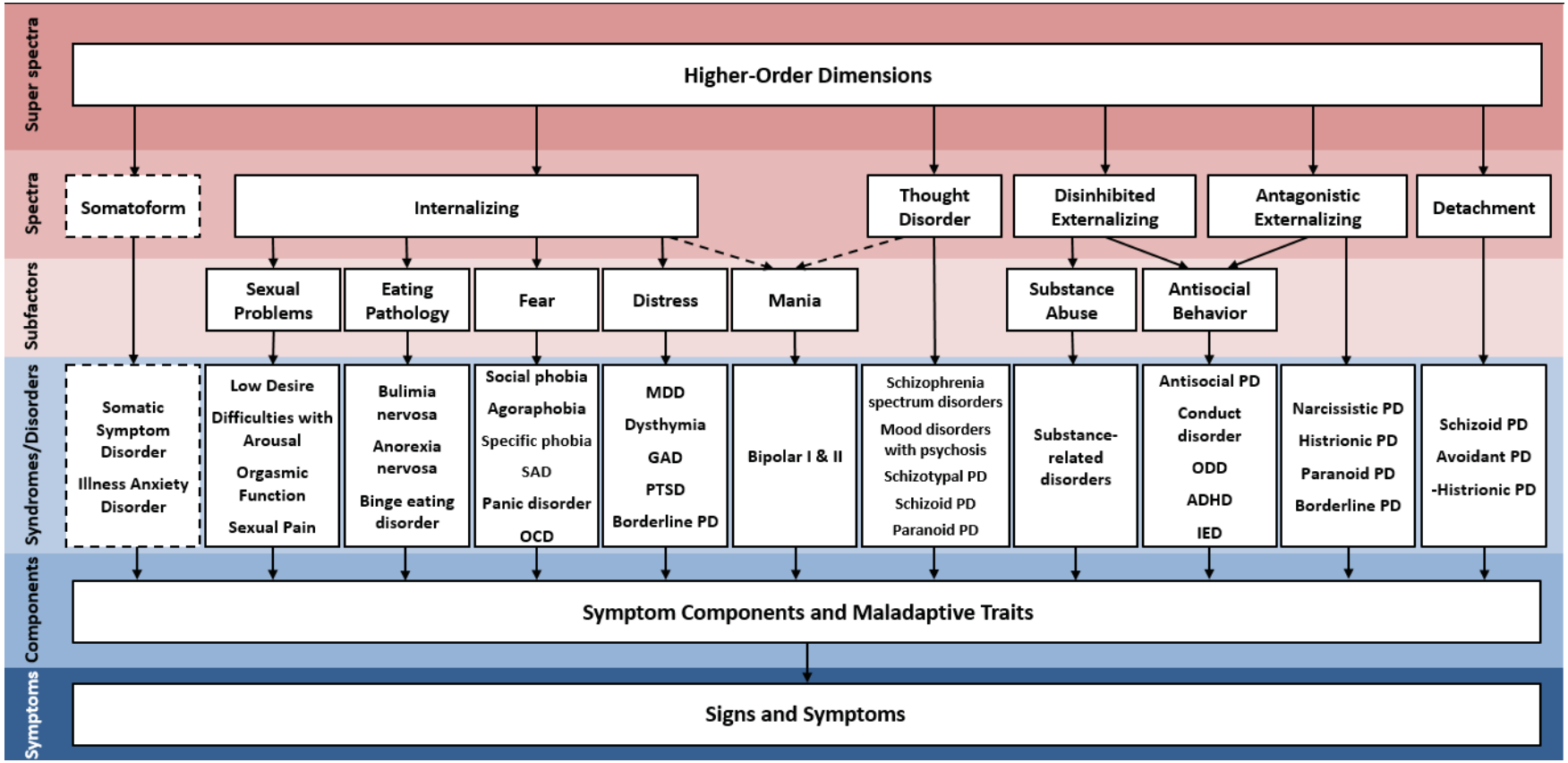
Comparison of Fit Statistics Between Competing eHiTOP and eRDoC GMM Models

Fit Statistic	eHiTOP Spectra GMM++	eHiTOP Subfactor GMM	eRDoC 6 Domain GMM++	eRDoC 5 Domain GMM
χ^2, (df; p-value)	480.344 (215; 0.000)	576.04 (285; 0.000)	503.18 (215; 0.000)	441.15 (201; 0.000)
χ^2/df ratio	2.23	2.02	2.34	2.19
RMSEA (95% CI)	0.062 (0.055 - 0.069)*	0.056 (0.050 - 0.063)*	0.065 (0.057 - 0.072)*	0.061 (0.053 – 0.069)*
SRMR	0.07*	0.06*	0.065*	0.062*
CFI	0.935	0.94	0.932	0.941
TLI	0.94*	0.94*	0.932*	0.941*
AIC	7769.73	7777.42	7764.34	26204.73
BIC	7894.293	7939.73	7888.90	26321.75

Note: ++ Indicates GMM models included in the final analyses. * Indicates adequate fit according to rules of thumb (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006). ** Indicates good fit according to rules of thumb (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006).

Figure 1

Spectra of the Hierarchical Taxonomy of Psychopathology



Note: Dashed lines indicate elements of the model that were included on provisional basis and require more study. Disorders with most prominent cross-loadings are listed in multiple places. Minus sign indicates negative association between histrionic personality and detachment spectrum. Kotov et al., 2017, p. 462.

Figure 2

GMM Model with PDSQ Scores Predicting OQ Total Score Trajectory Over Time

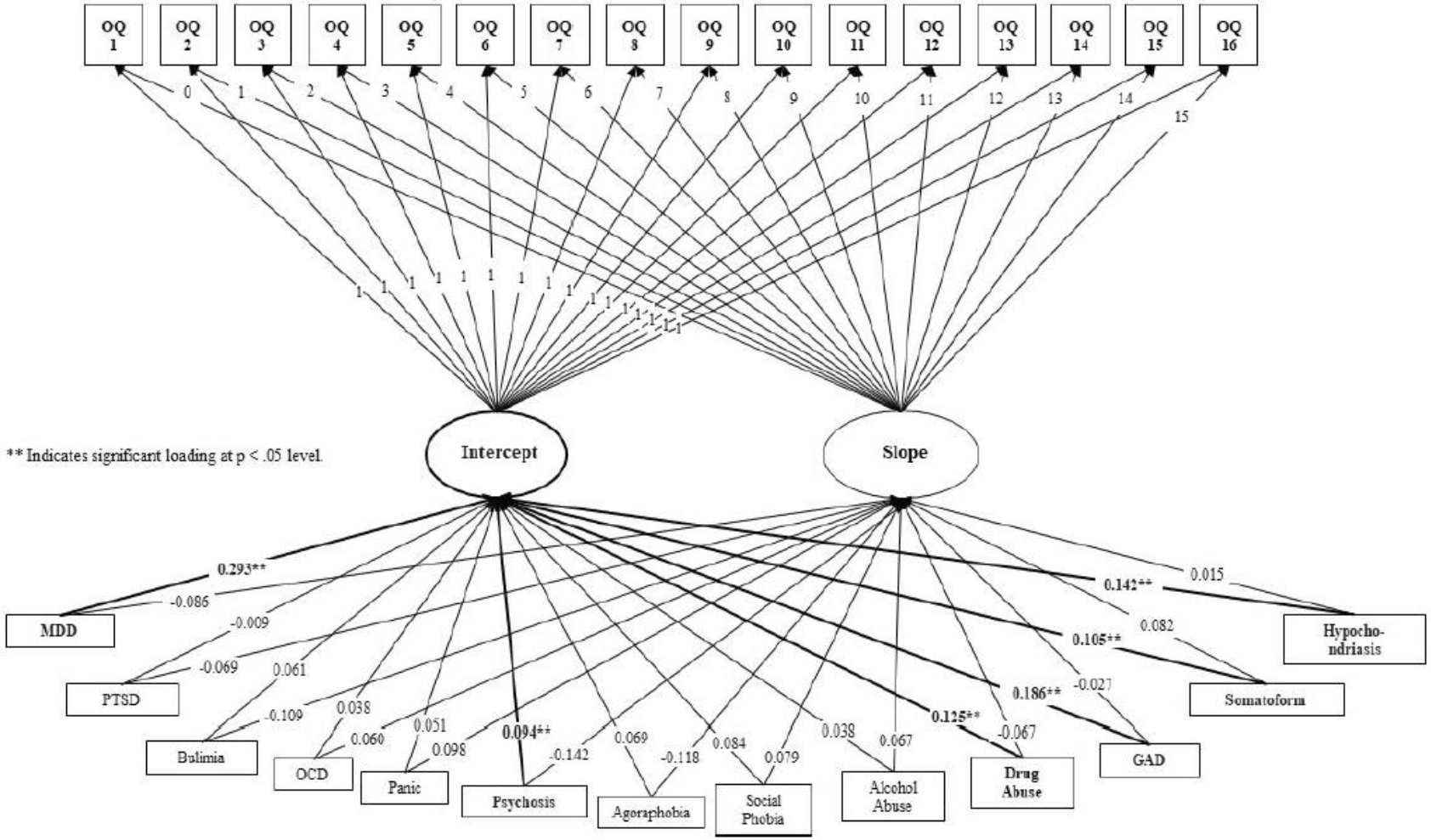


Figure 3

*GMM Model with eRDoC Scores Predicting OQ Total Score Trajectory Over Time**

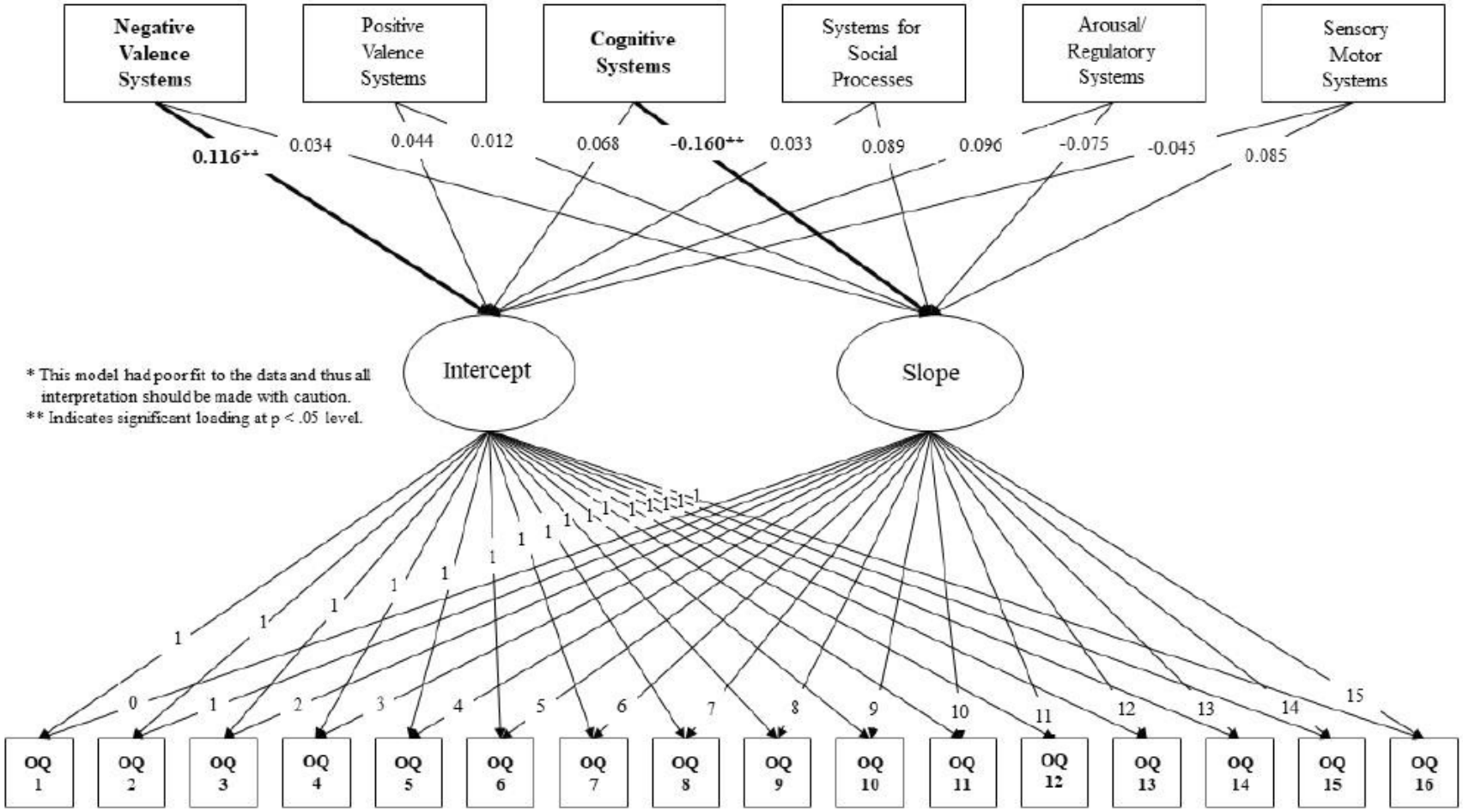
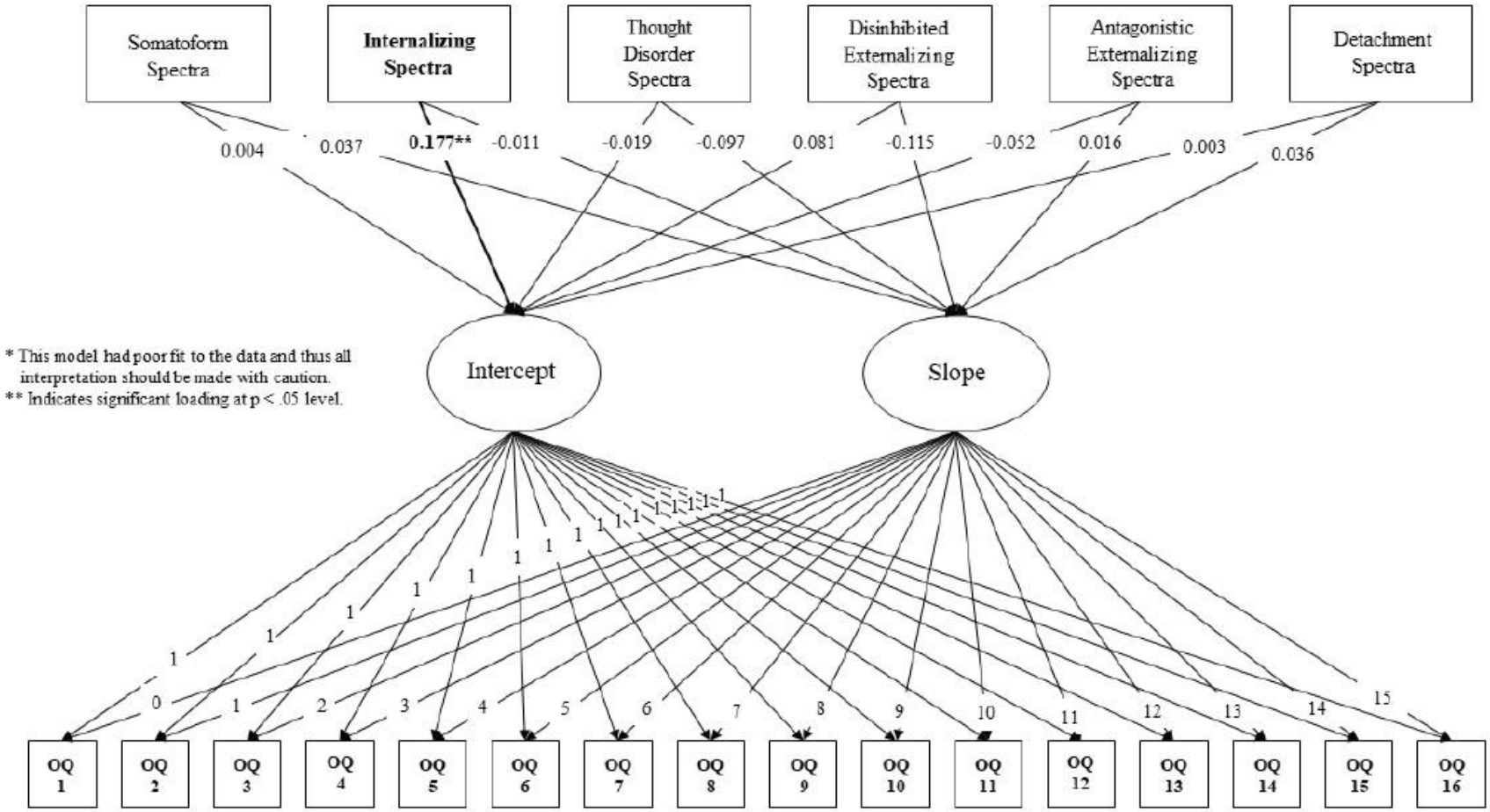


Figure 4

GMM Model with eHiTOP Scores Predicting OQ Total Score Trajectory Over Time



REFERENCES

Achenbach, T. M. (2009). *The Achenbach System of Empirically Based Assessment (ASEBA): Development, findings, theory, and applications*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families

Achenbach, T. M., Howell, C. T., Quay, H. C., Conners, C. K., & Bates, J. E. (1991). National survey of problems and competencies among four-to sixteen-year-olds: parents' reports for normative and clinical samples. *Monographs of the Society for Research in Child Development, 56*, i-130. Doi: 10.2307/1166156

Ambrosini, P. J., Bennett, D. S., Cleland, C. M., & Haslam, N. (2002). Taxonicity of adolescent melancholia: a categorical or dimensional construct?. *Journal of Psychiatric Research*, *36*, 247-256. doi: 10.1016/S0022-3956(02)00011-0

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Anderson, J., Snider, S., Sellbom, M., Krueger, R., & Hopwood, C. (2014). A comparison of the DSM-5 Section II and Section III personality disorder structures. *Psychiatry Research*, *216*, 363-372. doi: 10.1016/j.psychres.2014.01.007

Andrews, G., Brugha, T., Thase, M., Duffy, F. F., Rucci, P., & Slade, T. (2008). Dimensionality and the category of major depressive episode. In J. E., Helzer, H. C. Kramer, R. F. Krueger, H-U. Wittchen, P. J. Sirovatika, & D. A. Regier (Eds.), *Dimensional approaches to diagnostic classification. Refining the research agenda for DSM-V* (pp. 35-51). Washington, DC: American Psychiatric Association.

Baldwin, G., & Shean, G. D. (2006). A taxometric study of the Center for Epidemiological Studies depression scale. *Genetic, Social, and General Psychology Monographs, 132,* 101-128. Doi: 10.3200/MONO.132.2.101-128

Beauchaine, T. P., & Hinshaw, S. P. (Eds.). (2017). *Child and Adolescent Psychopathology*. John Wiley & Sons.

Biering, K., Hjollund, N. K., and Frydenberg, M. (2015). Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clinical Epidemiology, 7,* 91-106. doi: 10.2147/CLEP.S72247

Bijl, R. V., Ravelli, A., & van Zessen, G. (1998). Prevalence of psychiatric disorders in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology, 33*, 587-595. doi: 10.1007/s001270050098

Blanco, C., Wall, M. M., He, J. P., Krueger, R. F., Olfson, M., Jin, C. J., ... & Merikangas, K. R. (2015). The space of common psychiatric disorders in adolescents: comorbidity structure and individual latent liabilities. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*, 45-52. doi: 10.1016/j.jaac.2014.10.007

Boswell, D. L., White, J. K., Sims, W. D., Harrist, R. S., & Romans, J. S. (2013). Reliability and validity of the Outcome Questionnaire–45.2. *Psychological Reports*, *112*, 689-693. doi: 10.2466/02.08.PR0.112.3.689-693

Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, *11*, 515-524. doi: 10.1016/0191-8869(90)90065-Y

Boyd, J. H., Burke, J. D. J., Gruenberg, E., Holzer, C. E. R., Rae, D. S., George, L. K., et al. (1984). Exclusion criteria of DSM-III: A study of co-occurrence of hierarchy-free syndromes. *Archives of General Psychiatry, 41,* 983–989. doi: 10.1001/archpsyc.1984.01790210065008

Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. *Psychological Assessment, 21*, 256-271. doi: 10.1037/a0016608

Carragher, N., Krueger, R. F., Eaton, N. R., Markon, K. E., Keyes, K. M., Blanco, C., ... & Hasin, D. S. (2014). ADHD and the externalizing spectrum: direct comparison of categorical, continuous, and hybrid models of liability in a nationally representative sample. *Social Psychiatry and Psychiatric Epidemiology, 49*, 1307-1317. doi: 10.1007/s00127-013-0770-3

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... & Moffitt, T. E. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders?. *Clinical Psychological Science*, *2*, 119-137. doi: 10.1177/2167702613497473

Cassin, S. E., & von Ranson, K. M. (2005). Personality and eating disorders: a decade in review. *Clinical Psychology Review*, *25*, 895-916. doi: 10.1016/j.cpr.2005.04.012

Chinman, M. J., Allende, M., Weingarten, R., Steiner, J., Tworkowski, S., & Davidson, L. (1999). On the Road to Collaborative Treatment Planning: Consumer and Provider Perspectives. *The Journal of Behavioral Health Services & Research, 26*, 211-218. doi: 0.1007/BF02287492

Clark, L. A. (2007). Assessment and diagnosis of personality disorder: perennial issues and emerging conceptualization. *Annual Review of Psychology, 58,* 227-258. doi: 10.1146/annurev.psych.57.102904.190200

Clark, L. A., & Watson, D. (2006). Distress and fear disorders: an alternative empirically based taxonomy of the 'mood' and 'anxiety' disorders. *The British Journal of Psychiatry, 189*, 481-483. doi: 10.119 2/bjp.bp.106.03825

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46*, 121-153. doi: 10.1146/annurev.ps.46.020195.001005

Commonwealth of Massachusetts MassHealth (2014) *Provider Manual Series: Mental Health Center Manual*. 130 CMR 429.000 Retrieved from https://www.mass.gov/doc/mental-health-center-regulations-0/download

Conway, C. C., & Brown, T. A. (2018). Evaluating dimensional models of psychopathology in outpatients diagnosed with emotional disorders: A cautionary tale. *Depression and Anxiety*, *35*, 898-902. doi: 10.1002/da.22740

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., ... & Eaton, N. R. (2019). A hierarchical taxonomy of psychopathology can transform mental health research. *Perspectives on Psychological Science*, *14*, 419-436. doi: 10.1177/1745691618810696

Corruble, E., Legrnad, J. M., Zvenigorowski, H., Duret, C., & Guelfi, J. D., Concordance between self-report and clinician's assessment of depression. *Journal of Psychiatric Research, 5,* 457-465. doi: 10.1016/S0022-3956(99)00011-4

Costa, P. T., & MacCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual*. Psychological Assessment Resources.

Eisenberg, N., Valiente, C., Spinrad, T. L., Cumberland, A., Liew, J., Reiser, M., ... & Losoya, S. H. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology*, *45*, 988-1008. doi: 10.1037/a0016213

Eubanks, C. F. & Hunter, E. B. (2020) HiTOP and Psychotherapy Integration: Promising Potential. *Journal of Psychotherapy Integration, 30*, 498-505. doi: 10.1037/int0000254

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A. Jr., Winokur, G., & Munoz, R. (1972) Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26,* 57-63. doi: 10.1001/archpsyc.1972.01750190059011

First, M. B., Bell, C. B., Cuthbert, B., Krystal, J. H., Malison, R., Offord, D. R., … Wisner, K. L. (2002). Personality disorders and relational disorders: A research agenda for addressing crucial gaps in DSM. In D. J. Kupfer, M. B. First, & D. A. Reiger (Eds.), *A research agenda for DSM-V* (123-199). Washington, DC: American Psychiatric Association.

Forbes, M. K., Tackett, J. L., Markon, K. E., & Krueger, R. F. (2016). Beyond comorbidity: Toward a dimensional and hierarchical approach to understanding psychopathology

across the life span. *Development and Psychopathology*, *28*, 971-986. doi: 10.1017/S0954579416000651

Fowler, J. C., Sharp, C., Kalpakci, A., Madan, A., Clapp, J., Allen, J. G., … & Oldham, J. M. (2015). A dimensional approach to assessing personality functioning: Examining personality trait domains utilizing *DSM-IV* personality disorder criteria. *Comprehensive Psychiatry,56*, 75-84. doi: 10.1016/j.comppsych.2014.09.001

Frances, A. (Ed.). (1995). *Task Force DSM-IV*. American Psychiatric Association.

Franklin, C. L., Strong, D. R., & Greene, R. L. (2002). A taxometric analysis of the MMPI-2 depression scales. *Journal of Personality Assessment*, *79*, 110-121. doi: 10.1207/S15327752JPA7901_07

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of Personality and Social Psychology, 59*, 1216-1229. doi: 10.1037/0022-3514.59.6.1216

Gorka, S.M., Lieberman, L., Shankman, S.A., Phan, K.L., (2017). Startle potentiation to uncertain threat as a psychophysiological indicator of fear-based psychopathology: An examination across multiple internalizing disorders. *Journal of Abnormal Psychology, 126,* 8-18. doi: 10.1037/abn0000233

Grey, N., Kennedy, P. (1993). The functional independence measure: a comparative study of clinician and self ratings. *Paraplegia, 31*. 457-461. doi: 10.1038/sc.1993.74

Hankin, B. L., Fraley, R. C., Lahey, B. B., & Waldman, I. D. (2005). Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *Journal of Abnormal Psychology*, *114*, 96. Doi: 10.1037/0021-843X.114.1.96

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychological Medicine, 42*, 903-920. doi:10.1017/S0033291711001966

Hasler, G., Drevets, W. C., Manji, H. K., & Charney, D. S. (2004). Discovering endophenotypes for major depression. *Neuropsychopharmacology, 29*, 1765-1781. doi: 10.1038/sj.npp.1300506

Helzer, J. E., Wittchen, H. U., Krueger, R. F., & Kraemer, H. C. (2008). Dimensional options for DSM-V: The way forward. In J. E. Helzer, H. C. Kraemer, R. F. Hrueger, H. Wittechen, P. J. Sirovatika, & D. A. Regier (Eds.), *Dimensional approaches to diagnostic classification: Refining the research agenda for DSM-V* (115-127). Arlington, VA: American Psychiatric Association

Hopwood, C. J., Bagby, R. M., Gralnick, T., Ro, E., Ruggero, C., Mullins-Sweatt, S., ... & Zimmermann, J. (2020). Integrating psychotherapy with the hierarchical taxonomy of

psychopathology (HiTOP). *Journal of Psychotherapy Integration, 30*, 477-497. doi: 10.1037/int0000156

Hopwood, C. J., Kotov, R., Krueger, R. F., Watson, D., Widiger, T. A., Althoff, E. B., … & Zimmerman, J. (2017). The time has come for dimensional personality disorder diagnosis. *Personality and Mental Health, 12,* 82-86. doi: 10.1002/pmh.1408

Iacono, W. G., Malone, S. M., & McGue, M. (2003). Substance use disorders, externalizing psychopathology, and P300 event-related potential amplitude. *International Journal of Psychophysiology*, *48*, 147-178. doi: 10.1016/S0167-8760(03)00052-7

Insel, T. R. (2014). The NIMH research domain criteria (RDoC) project: Precision medicine for psychiatry. *American Journal of Psychiatry, 4*, 395-397. doi: 10.1176/appi.ajp.2014.14020138

Jobes, D. A. (2012). The Collaborative Assessment and Management of Suicidality (CAMS): An evolving evidence-based clinical approach to suicidal risk. *The American Association of Suicidology, 42,* 640-653. doi: 10.1111/j.1943-278X.2012.00119.x

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development*, *65*, 160-178. doi: 10.1111/j.1467-8624.1994.tb00742.x

Kaufman, J., Gelernter, J., Hudziak, J. J., Tyrka, A. R., & Coplan, J. D. (2015). The research domain criteria (RDoC) project and studies of risk and resilience in maltreated children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 617-625. doi: 10.1016/j.jaac.2015.06.001

Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. *The British Journal of Psychiatry*, *127*, 305-315. Doi: 10.1192/bjp.127.4.305

Kendell, R. E. (1989). Clinical validity. *Psychological Medicine*, *19*, 45-55. doi: 10.1017/S0033291700011016

Kessler, R. C., Zhao, S., Blazer, D. G., & Swartz, M. (1997). Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. *Journal of Affective Disorders*, *45*, 19-30. doi: 10.1016/S0165-0327(97)00056-6

Kotov, R., Chang, S. W., Fochtmann, L. J., Mojtabai, R., Carlson, G. A., Sedler, M. J., & Bromet, E. J. (2010). Schizophrenia in the internalizing-externalizing framework: a third dimension?. *Schizophrenia Bulletin*, *37*, 1168-1178. doi: 10.1093/schbul/sbq024

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychological Bulletin*, *136*, 768-821. doi: 10.1037/a0020327

Kotov, R., Krueger, R. F., & Watson, D. (2018). A paradigm shift in psychiatric classification: the Hierarchical Taxonomy Of Psychopathology (HiTOP). *World Psychiatry*, *17*, 24-25. doi: 10.1002/wps.20478

Kotov R., Jonas K. G., Carpenter W. T., Dretsch M. N., Eaton N. R., … Watson, D. (2020). Validity and utility of hierarchical taxonomy of psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry,* 19, 151–172. doi: 10.1002/wps.20730

Kotov, R., Krueger, R. F., Watson, D., Cicero, D. C., Conway, C. C., DeYoung, C. G., ... & Wright, A. G. (2021). The hierarchical taxonomy of psychopathology (HiTOP): A quantitative nosology based on consensus of evidence. *Annual Review of Clinical Psychology*, *17*, 83-108. doi: 10.1146/annurev-clinpsy-081219-093304

Kotov, R., Moffitt, T. E., Krueger, R. F., Morey, L. C., Watson, D., Millins-Sweatt, S. N., … & Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*, 454- 477. doi: 10.1037/abn0000258

Kotov, R., Perlman, G., Gámez, W., & Watson, D. (2015). The structure and short-term stability of the emotional disorders: A dimensional approach. *Psychological Medicine, 45*, 1687-1698. doi: 10.1017/S0033291714002815

Kotov, R., Ruggero, C. J., Krueger, R. F., Watson, D., Yuan, Q., & Zimmerman, M. (2011). New dimensions in the quantitative classification of mental illness. *Archives of General Psychiatry*, *68*, 1003-1011. doi: 10.1001/archgenpsychiatry.2011.107

Krueger, R. F., Markon, K. E., Patrick, C. J., and Iacono, W. G. (2005). Externalizing Psychopathology in Adulthood: A Dimensional-Spectrum Conceptualization and Its Implications for *DSM–V*. *Journal of Abnormal Psychology, 114*, 537-550. doi: 10.1037/0021-843X.114.4.537

Krueger, R. F., Chentsova-Dutton, Y. E., Markon, K. E., Goldberg, D. P., & Ormel, J. (2003). A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *Journal of Abnormal Psychology, 107*, 216-227. doi: 10.1037/0021-843X.112.3.437

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine*, *42*, 1879-1890. doi: 10.1017/S0033291711002674

Krueger, R. F., & Markon, K. E. (2011). A dimensional-spectrum model of psychopathology: progress and opportunities. *Archives of General Psychiatry, 68*, 10-11. doi: 10.1001/archgenpsychiatry.2010.188

Krueger, R. F., & Markon, K. E. (2014). The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology*, *10*, 477-501. doi: 10.1146/annurev-clinpsy-032813-153732

Krueger R. F., Kotov R., Watson D., Forbes M. K., Eaton N .R., … Zimmerman, J. (2018). Progress in achieving quantitative classification of psychopathology. *World Psychiatry, 17,* 282–293. doi: 10.1002/wps.20566

Lahey, B. B., Kruger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017) Validity and utility of the general factor of psychopathology. *World Psychiatry, 16*, 142-143. doi: 10.1002/wps.20410

Latzman, R. D., DeYoung, C. G., & The Hitop Neurobiological Foundations Workgroup. *et al.* (2020). Using empirically-derived dimensional phenotypes to accelerate clinical neuroscience: The hierarchical taxonomy of psychopathology (HiTOP) framework. *Neuropsychopharmacology*, *45*, 1083-1085. doi: 10.1038/s41386-020-0639-6

Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. C., Shimokawa, K., Christopherson, C., & Burlingame, G. M. (2004). *Administration and Scoring Manual for the OQ-45.2 (Outcome Questionnaire)*. Salt Lake City, UT: OQ Measures, LLC

Lengua, L. J., West, S. G., & Sandler, I. N. (1998). Temperament as a predictor of symptomatology in children: Addressing contamination of measures. *Child Development*, *69*, 164-181. doi: 0009-3920/98/6901-0005$01.00

Lengua, L. J. (2006). Growth in temperament and parenting as predictors of adjustment during children's transition to adolescence. *Developmental Psychology*, *42*, 819-832. doi: 10.1037/0012-1649.42.5.819

Lewis, D. D., Yang, Y., Rose, T. G., & Li, F. (2004). RCVI: A new benchmark collection for text categorization research. *Journal of Machine Learning Research, 5*, 361-397. retrieived from: https://jmlr.csail.mit.edu/papers/volume5/lewis04a/lewis04a.pdf

Lieblich, S. M., Castle, D. J., Pantelis, C., Hopwood, M., Young, A. H., & Everall, I. P. (2015). High heterogeneity and low reliability in the diagnosis of major depression will impair the development of new drugs. *BJPsych Open*, *1*, e5-e7. doi: 10.1192/bjpo.bp.115.000786

Lilienfeld, S. O., Waldman, I. D., & Israel, A. C. (1994). A critical examination of the term and concept of 'comorbidity' in psychopathology research. *Clinical Psychology Science and Practice, 1*, 71-83. doi: 10.1111/j.1468-2850.1994.tb00007.x

Little, T. D., Jorgenson, T. D., Lang, K. M., and Moore, E. W. G. (2014) On the joys of missing data. *Journal of Pediatric Psychology, 39*, 151-162. doi: 10.1093/jpepsy/jst048

Livesley, W. J. (2003). Diagnostic dilemmas in classifying personality disorder. In K. A. Phillips, M. B. First, & H. A. Pincus (Eds.), *Advancing DSM: Dilemmas in psychiatric diagnosis* (pp. 153-189). Arlington, VA: American Psychiatric Association.

Love, P. K. (2018). *Clinical utility of research domain criteria* (Unpublished Master's thesis). University of North Texas, Denton, TX, US.

Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: a meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, *27*, 101-114. doi: 10.1007/s10862-005-5384-y

Markon, K. E., & Krueger, R. F. (2005). Categorical and continuous models of liability to externalizing disorders: A direct comparison in NESARC. *Archives of General Psychiatry, 62*, 1352-1359. doi: 10.1001/archpsyc.62.12.1352

Markon, K. E. (2010). Modeling psychopathology structure: A symptom-level analysis of Axis I and II disorders. *Psychological Medicine, 40*, 273-288. doi: 10.1017/S0033291709990183

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139-175. doi: 10.1037/0022-3514.88.1.139

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., ... & Rohde, L. A. (2017). A general psychopathology factor (P factor) in children: structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology*, *126*, 137-148. doi: 10.1037/abn0000205

McCoy, T. H., Castro, V. M., Rosenfield, H. R., Cagan, A., Kohanne, I. S., Perlis, R. H. (2015). A clinical perspective on the relevance of research domain criteria in electronic health records. *American Journal of Psychiatry, 172*, 316-320. doi: 10.1176/appi.ajp.2014.14091177.

McCoy Jr, T. H., Yu, S., Hart, K. L., Castro, V. M., Brown, H. E., Rosenquist, J. N., ... & Perlis, R. H. (2018). High throughput phenotyping for dimensional psychopathology in electronic health records. *Biological Psychiatry*, *83*(12), 997-1004. doi: 10.1016/j.biopsych.2018.01.011

McElroy E, Belsky J, Carragher N, Fearon P, Patalay P. 2018. Developmental stability of general and specific factors of psychopathology from early childhood to adolescence: dynamic mutualism or p-differentiation? *Journal of Child Psychology and Psychiatry*, 59, 667–675. doi:10.1111/jcpp.12849

McNulty, J. L., & Overstreet, S. R. (2014). Viewing the MMPI–2–RF structure through the Personality Psychopathology Five (PSY–5) lens. *Journal of Personality Assessment*, *96*, 151-157. doi: 10.1080/00223891.2013.840305

Meehl, P. E., & Golden, R. (1982). Taxometirc methods. In Kendall, P. C. & Butcher, J. N. (Eds.), *Handbook of Research Methods in Clinical Psychology* (127-181). New York, NY: Wiley.

Michelini G., Palumbo I. M., DeYoung C. G., Latzman R., & Kotov R. (2020). Linking RDoC and HiTOP: a new interface for advancing psychiatric nosology and neuroscience. *Clinical Psychology Review, 86*, 102025. doi: /10.1016/j.cpr.2021.102025

Morey, L. C., Hopwood, C. J., Gunderson, J. G., Skodol, A. E., Shea, M. T., Yen, S., ... & McGlashan, T. H. (2007). Comparison of alternative models for personality disorders. *Psychological Medicine*, *37*, 983-994. doi: 10.1017/S0033291706009482

Morey, L. C., Hopwood, C. J., Markowitz, J. C., Gunderson, J. G., Grilo, C. M., McGlashan, T. H., ... & Skodol, A. E. (2012). Comparison of alternative models for personality disorders, II: 6-, 8-, and 10-year follow-up. *Psychological Medicine*, *42*(8), 1705-1713. doi:10.1017/S0033291711002601.

Muthén, B. (2001a). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Collins, L.M. & Sayer, A. (eds.), *New Methods for the Analysis of Change* (pp. 291-322). Washington, D.C.: APA.

Muthén, B. (2001b). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (eds.), *New Developments and Techniques in Structural Equation Modeling* (pp. 1-33). Lawrence Erlbaum Associates.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), *Handbook of Quantitative Methodology for the Social Sciences* (pp. 345-368). Newbury Park, CA: Sage Publications.

Muthén, B., Brown, C.H., Masyn, K., Jo, B., Khoo, S.T., Yang, C.C., Wang, C.P., Kellam, S., Carlin, J. & Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics, 3*, 459-475. doi: 10.1093/biostatistics/3.4.459

Muthén, B. & Masyn, K. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics*, *30*, 27-58. doi: 10.3102/10769986030001027

Muthén, B. & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55*, 463-469. doi: 10.1111/j.0006-341X.1999.00463.x

National Institutes of Mental Health (2013). *Development and Definitions of the RDoC Domains and Constructs*. Retrieved from: https://www.nimh.nih.gov/research-priorities/rdoc/development-and-definitions-of-the-rdoc-domains-and-constructs.shtml

National Institutes of Mental Health (2021). *RDoC Matrix.* Retrieved from: https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/constructs/rdoc-matrix.shtml

Ormel, J., Raven, D., van Oort, F., Hartman, C. A., Reijneveld, S. A., Veenstra, R., ... & Oldehinkel, A. J. (2015). Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. *Psychological Medicine, 45*, 345-360. doi:10.1017/S0033291714001469

Perkins, E. R., Joyner, K. J., Patrick, C. J., Bartholow, B. D., Latzman, R. D., DeYoung, C. G., ... & Zald, D. H. (2020). Neurobiology and the hierarchical taxonomy of psychopathology:

progress toward ontogenetically informed and clinically useful nosology. *Dialogues in Clinical Neuroscience*, *22*, 51-63. doi: 10.31887/DCNS.2020.22.1/eperkins

Phillips, K. A., Price, L. H., Greenberg, B. D., & Rasmussen, S. A. (2003). Should the DSM diagnostic groupings be changed? In K. A. Phillips, M. B. First, & H. A. Pincus (Eds.), *Advancing DSM: Dilemmas in psychiatric diagnosis* (pp. 57-84). Arlington, VA: American Psychiatric Association.

Pincus, H. A., Frances, A., Davis, W. W., First, M. B., & Widiger, T. A. (1992). DSM-IV and new diagnostic categories: holding the line on proliferation. *American Journal of Psychiatry*, *149*, 112-117. doi: 10.1176/ajp.149.1.112

Pincus, H. A., McQueen, L. E., & Elinson, L. (2003). Subthreshold mental disorders: Nosological and research recommendations. In K. A. Phillips, M. B. First, & H. A. Pincus (Eds.), *Advancing DSM: Dilemmas in psychiatric diagnosis* (pp. 129-144). Arlington, VA, US: American Psychiatric Association.

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 Field Trials in the United States and Canada, part II: test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry, 168*, 1186-1194. doi: 10.1177/1073191113486183

Rodriguez-Seijas, C., Eaton, N. R., & Krueger, R. F. (2015). How transdiagnostic factors of personality and psychopathology can inform clinical assessment and intervention. *Journal of Personality Assessment*, *97*, 425-435. doi: 10.1080/00223891.201

Ross, S. R., Lutz, C. J., & Bailley, S. E. (2002). Positive and negative symptoms of schizotypy and the five-factor model: a domain and facet level analysis. *Journal of Personality Assessment*, *79*, 53-72. doi: 10.1207/S15327752JPA7901_04

Røysamb, E., Kendler, K. S., Tambs, K., Orstavik, R. E., Neale, M. C., Aggen, S. H. … Reichbørn-Kjennerud, T. (2011). The joint structure of DSM-IV Axis I and Axis II disorders. *Journal of Abnormal Psychology, 120*, 198-209. doi: 10.1037/a0021660

Rubin, K. H., & Mills, R. S. (1991). Conceptualizing developmental pathways to internalizing disorders in childhood. *Canadian Journal of Behavioural Science*, *23*, 300-317. doi: 10.1037/h0079019

Ruggero, C. J., Callahan, J. L., Dornbach-Bender, A., Tackett, J. L., & Kotov, R. (2019). Disquiet in nosology: A primer on an emerging, empirically-based approach to classifying mental illness and implications for training. *The Behavior Therapist, 42,* 208-211.

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E., ... & Zimmermann, J. (2019). Integrating the hierarchical taxonomy of psychopathology (HiTOP) into clinical practice. *Journal of Consulting and Clinical Psychology, 87*, 1069-1084. doi:10.1037/ccp0000452

Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology*, *109*, 473-487. doi: 10.1037/0021-843X.109.3.473

Sakashita, C., Slade, T., & Andrews, G. (2007). An empirical investigation of two assumptions in the diagnosis of DSM-IV major depressive episode. *Australian and New Zealand Journal of Psychiatry, 41*, 17-23. doi: 10.1080/00048670601050440

Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review, 28*, 1326-1342. doi: 10.1016/j.cpr.2008.07.002

Schermelleh-Engel, K. & Moosbrugger, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8,* 23-74.

Sellbom, M. (2017). Mapping the MMPI–2–RF Specific Problems scales onto extant psychopathology structures. *Journal of Personality Assessment*, *99*, 341-350. doi: 10.1080/00223891.2016.1206909

Shields, A. N., Giljen, M., España. R. A., & Tackett, J. L. (2021). The p factor and dimensional structural models of youth personality pathology and psychopathology. *Current Opinion in Psychology*, *37*, 21-25. doi: /10.1016/j.copsyc.2020.06.005

Skeem, J. L., Polaschek, D. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest, 12*, 95-162. doi: 10.1177/1529100611426706

Slade, T., & Andrews, G. (2005). Latent structure of depression in a community sample: a taxometric analysis. *Psychological Medicine*, *35*, 489-497. Doi: 10.1017/S0033291704003708

Slade, T. (2007). The descriptive epidemiology of internalizing and externalizing psychiatric dimensions. *Social Psychiatry and Psychiatric Epidemiology, 42*, 554-560. doi: 10.1007/s00127-007-0200-5

Stanhope, V., Ingoglia, C., Schmelter, B., & Marcus, S. C. (2013). Impact of person-centered planning and collaborative documentation on treatment adherence. *Psychiatric Services, 64*, 76-79. doi: 10.1176/appi.ps.201100489

Substance Abuse and Mental Health Service Administration (2015). *Certified Community Behavioral Health Clinic Certification Guide*. SAMASA: Rockville, MD Retrieved on May 8. 2021 from: https://www.samhsa.gov/sites/default/files/programs_campaigns/state-certification-guide.pdf

Sunderland M. & Slade, T. (2015) The relationship between internalizing psychopathology and suicidality, treatment seeking, and disability in the Australian population. *Journal of Affective Disorders*, *171*, 6-12. doi: 10.1016/j.jad.2014.09.012

Tackett, J. L., Silberschmidt, A. L., Krueger, R. F., Sponheim, S. R. (2008) A dimensional model of personality disorder: Incorporating DSM Cluster A characteristics. *Journal of Abnormal Psychology, 117,* 454-459. doi: 10.1037/0021-843X.117.2.454

Taylor, G. J., & Bagby, R. M. (2012). The alexithymia personality dimension. In T. A. Widiger (Ed.), *Oxford library of psychology. The Oxford handbook of personality disorders* (pp. 648-673). New York, NY, US: Oxford University Press. doi: 10.1093/oxfordhb/9780199735013.013.0030

Teesson, M., Slade, T., & Mills, K. (2009). Comorbidity in Australia: Findings of the 2007 National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry, 43*, 606-614. doi: 10.1080/00048670902970908

Tomar, S. S. (2017). Text mining in R: A tutorial. Retrieved from: https://www.springboard.com/blog/text-mining-in-r/

Trull, T. J., & Durrett, C. A. (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology, 1*, 355-380. doi: 10.1037/0003-066X.62.2.71

Trull, T. J., & Sher, K. J. (1994). Relationship between the five-factor model of personality and Axis I disorders in a nonclinical sample. *Journal of Abnormal Psychology*, *103*, 350-360. 10.1037/0021-843X.103.2.350

Uher, R., Perlis, R. H., Placentino, A., Dernovšek, M. Z., Henigsberg, N., Mors, O., ... & Farmer, A. (2012). Self-report and clinician-rated measures of depression severity: can one replace the other?. *Depression and Anxiety*, *29*, 1043-1049. doi: 10.1002/da.21993

Uliaszek, A. A., Al-Dajani, N., & Bagby, R. M. (2015). The relationship between psychopathology and a hierarchical model of normal personality traits: Evidence from a psychiatric patient sample. *Journal of Personality Disorders*, *29*, 719-734. doi: 10.1521/pedi_2014_28_170

Umphress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment*, *15*, 40-55. doi: 10.1177/073428299701500104

Üstün, T. B., & Sartorius, N. (Eds.). (1995). *Mental illness in general health care: an international study*. New York, NY: John Wiley & Sons.

Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome Questionnaire: Is it sensitive to changes in counseling center clients?. *Journal of Counseling Psychology*, *51*, 38-49. doi: 10.1037/0022-0167.51.1.38

Vermeerch, D. A., Lambert, M. J., & Burlingame, G. M. (2010). Outcome Questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74*, 242-261. doi: 10.1207/S15327752JPA7402_6

Verona, E., Sachs-Ericsson, N., & Joiner Jr, T. E. (2004). Suicide attempts associated with externalizing psychopathology in an epidemiological sample. *American Journal of Psychiatry*, *161*, 444-451. doi: 10.1176/appi.ajp.161.3.444

Wakefield, J. C. (2016). Diagnostic issues and controversies in DSM-5: Return of the false positives problem. *Annual Review of Clinical Psychology, 12,* 105-132. doi: 10.1146/annurev-clinpsy-032814-112800

Waldman, I. D., Poore, H. E., van Hulle, C., Rathouz, P. J., & Lahey, B. B. (2016). External validity of a hierarchical dimensional model of child and adolescent psychopathology: tests using confirmatory factor analyses and multivariate behavior genetic analyses. *Journal of Abnormal Psychology*, *125*(8), 1053-1067. doi: 10.1037/abn0000183

Watson, D. (2003). Subtypes, specifiers, epicycles, and eccentrices: towards a more parsimonious taxonomy of psychopathology. *Clinical Psychology: Science and Practice, 10*, 233-238. doi:10.1093/clipsy/bpg013

Weston, R. & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34,* 719-751. doi: 10.1177/0011000006286345

Widiger, T. A., & Costa Jr, P. T. (2013). *Personality Disorders and the Five-Factor Model of Personality* (3rd ed.). Washington, DC, US: American Psychological Association. doi: 10.1037/13939-000

Widiger , T. A., & Crego, C. (2018) Mental disorders as discrete clinical conditions: Dimensional versus categorical classification. In D. C. Beidel & B. C. Frueh (eds.) *Adult psychopathology and diagnosis* (3-32). Hoboken, NJ: John Wiley & Sons, Inc.

Widger, T. A. & Edmundson, M. (2011). Diagnoses, dimensions, and DSM-V. In D. Richard & S. Huprich (Eds.), *Clinical psychology, assessment, treatment, and research* (69-88). Oxford, England: Elsevier.

Whiteside, S. P., & Lynam, D. R. (2009). Understanding the role of impulsivity and externalizing psychopathology in alcohol abuse: Application of the UPPS impulsive behavior scale. *Personality Disorders: Theory, Research, and Treatment*, *2,* 66-79. doi: 10.1037/1949-2715.S.1.69

Widiger, T. A., & Mullins-Sweatt, S. N. (2009). Five-factor model of personality disorder: a proposal for DSM-V. *Annual Review of Clinical Psychology, 5*, 197-220. doi: 10.1146/annurev.clinpsy.032408.153542

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: shifting to a dimensional model. *American Psychologist, 62,* 71-83. doi: 10.1037/0003-066X.62.2.71

Wildes, J. E., & Marcus, M. D. (2015). Application of the research domain criteria (RDoC) framework to eating disorders: Emerging concepts and research. *Current Psychiatry Reports*, *17*, 30. doi: 10.1007/s11920-015-0572-2

Wright, A. G. C., Krueger, R. F., Hobbs, M. J., Markon, K. E., Eaton, N. R., & Slade, T. (2013). The structure of psychopathology: Toward an expanded quantitative empirical model. *Journal of Abnormal Psychology, 122,* 281-294. doi: 10.1037/a0030133

Wygant, D. B., Sellbom, M., Sleep, C. E., Wall, T. D., Applegate, K. C., Krueger, R. F., & Patrick, C. J. (2016). Examining the *DSM-5* alternative personality disorder model operationalization of antisocial personality disorder and psychopathy in a male correctional sample. *Personality Disorders, Theory, Research, and Treatment, 7,* 229-239. doi: 10.1037/per0000179

Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164,* 557-565. doi: 10.1176/appi.ajp.164.4.557

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry, 56*, 29-34. doi: 0.1016/j.comppsych.2014.09.007

Zimmerman, M., & Mattia, J. I. (2001). A self-report scale to help make psychiatric diagnoses: The Psychiatric Diagnostic Screening Questionnaire. *Archives of General Psychiatry*, *58*, 787-794. Doi: 10.1001/archpsyc.58.8.787